A MORE GENERAL FORMULATION OF SEPARABLE LEAST SQUARES

V.G. Doví
E.P. Arato
L. Maga

Istituto di Scienze e Tecnologie dell'Ingegneria Chimica
Universitá di Genova
Via Opera Pia, 15
16145 Genova (Italy)

ABSTRACT. Separable least squares are generally written in the form

$$\| \underline{y} - \underline{\underline{A}}(\underline{g})\, \underline{c} \|^2 = \min \qquad (1)$$

where minimization has to be carried out with respect to the parameters $\underline{g}$ and $\underline{c}$. The latter enter linearly into the expression of the objective function and this leads to considerable simplifications. These simplifications are not possible if the functional to be minimized is of the form

$$\Sigma_j \| \underline{y}_j - \underline{\underline{A}}_j(\underline{g})\, \underline{c} \|^2 = \min \qquad (2)$$

In this communication we modify previous methods, so that they can deal with the more general case (2). We also show that the statistical properties of both linear and nonlinear estimates can be evaluated according to a procedure that enables us to keep the two classes of parameters separate. An important application is examined and its numerical solution worked out in detail.

Keywords. Data analysis; parameter estimation; multivariate analysis; separability properties; hierarchical optimization.

INTRODUCTION

In many optimization problems it is possible to take advantage of the special structure of the objective function and to reduce the number of the variables whith respect to which optimization is actually carried out.
Suppose the optimization is to be performed in the R^n space.
If it is possible to partitio the set of variables $\underline{x}$ in two different sets $\underline{x}_1 \in R^a$ and $\underline{x}_2 \in R^b$ (a+b=n) such that $\underline{x}_2$ optimum values of $\underline{x}_1$ can be obtained with little effort, it is convenient to perform the optimization with respect to the $\underline{x}_2$ only, optimum $\underline{x}_1$ being determined automatically at each iteration. Variables $\underline{x}_1$ and $\underline{x}_2$ are said to be separated.
The main advantage of separating variables are increased efficiency, i.e. a more rapid convergence, and robustness, i.e. improved convergence properties.
While reduction in the number of variable from n to b is a considerable advantage, we have to remember that each iteration requires a complete optimization with respect to the remaining a variables.
This is why the condition that for fixed $\underline{x}_2$ values optimization with respect to the $\underline{x}_1$ be quickly performed has to be satisfied for the method to be of practical interest.
Sometimes we are prepared to separate the variables even though the resulting algorithm is slower than the original one, provided the convergence properties are enhanced.
This generally occurs if some variables are strongly correlated and by means of separation it is possible to assign them to different sets $\underline{x}_1$ and $\underline{x}_2$ (Doví,1974).
In all other cases separation of variables is practically confined to the presence of linear variables in least squares problems of the type

$$\| \underline{y} - \underline{\underline{A}}(\underline{x}_2)\underline{x}_1 \|^2 = \min \qquad (3)$$

where the $\underline{y}$ are data and A is a matrix whose elements are arbitrary funtions of $\underline{x}_2$. The presence of arbitrary constraints on $\underline{x}_2$ does not alter the method in any substantial way. If the data are in the form of a one dimensional vector , i.e. if only the one variable y_μ is measured in each experiment, well established method and algorithmic implementations are avaible for the solution of the resulting problem (Golub,1973), (Ruhe,1980), (Golub,1986).
In this paper we generalize these methods to two-dimensional data sets, i.e. to all those cases in which more variables are measured in each experiment.
The resulting optimization problem is therefore of the type

$$\Sigma_j \| \underline{y}_j - \underline{\underline{A}}_j(\underline{g})\underline{c} \|^2 = \min \qquad (4)$$

Techniques to resolve problem (4) and to interpret the results statistically will be described in next two sections, whereas an important application will be examined in the following one.

ALGORITHM

System (3) can be solved like most least squares problems using either normal equa-

tions techniques or orthogonal decomposition. The former method is more efficient in that the number of operations required is half that of decomposition methods, but it requires precision ε^2 to obtain the same accuracy reached by decomposition algorithms whit precision ε.
Thus for fixed values of $\underline{x}_2$, $\underline{x}_1$ is given by

$$\underline{x}_1 = (\underline{\underline{A}}^*(\underline{x}_2)\underline{\underline{A}}(\underline{x}_2))^{-1}\ \underline{\underline{A}}^*(\underline{x}_2)\underline{Y} \qquad (5)$$

if normal equations are used and by

$$\underline{x}_1 = \underline{\underline{A}}^+(\underline{x}_2)\underline{Y} \qquad (6)$$

if pseudoinverse techniques are employed.
Of course differentiations of

$(\underline{\underline{A}}^*(\underline{x}_2)\ \underline{\underline{A}}(\underline{x}_2))^{-1}\ \underline{\underline{A}}^*(\underline{x}_2)$ and $\underline{\underline{A}}^+(\underline{x}_2)$ with

respect to $\underline{x}_2$ are required if gradient methods for the minimization of (3) are made use of.
The derivative of $(\underline{\underline{A}}^*(\underline{x}_1)\underline{\underline{A}}(\underline{x}_2))^{-1}\ \underline{\underline{A}}^*(\underline{x}_2)$

is generally computed by implicit differentiation (Provencher,1978) whereas the differentiation of the pseudoinverse A^+ is carried out using the Fréchet derivative of the ortogonal projector of A.
In this paper we examine the functional (4) and develop an algorithm based on generalized normal equations and capable of performing the minimization with respect to the non linear variables only. To this purpose we generalize the procedure first developped by Provencher (1978,1979) for the functional (3). To this purpose let us rewrite the functional (4) as follows

$$\Sigma_k\Sigma_i\ [\ y_{ik}-\Sigma_j x_{1j}A_{jki}(\underline{x}_2)\ ]^2 = \min \qquad (7)$$

For fixed values of $\underline{x}_2$ best values $\underline{x}_1$ are given by

$$x_{1j} = (A_{lik}A_{ikj})^{-1}A_{lik}y_{ik} \qquad (8)$$

where the tensor notation for indicating summation has been employed.
The inverse operation is supposed stabilized by suitable norm regularization if the matrix $(A^{ik}_{l}\ A^{j}_{ik})$ is not full rank.

Thus (7) reduces to

$$\Phi = \Sigma_k\Sigma_i\ [y_{ik}-A_{kij}(A_{l}A^{j}_{ikj})^{-1}A_{lik}y_{ik}]^2 \qquad (9)$$

and optimum values of the objective function Φ can be evaluated for every choice of $\underline{x}_2$. In this case too derivatives of Φ with respect to the non linear variables $\underline{x}_2$ contained in A are to be computed if gradient algorithms for the minimization of Φ are employed. On the other hand one generally has to employ this class of algorithms for otherwise convergence would be reached very slowly, if at all. In particular we are going to outline all the steps necessary to compute a typical Gauss-Newton iteration, in which the gradient of Φ is computed exactly and its hessian is approximated by omission of terms containing residuals. In order to compute the derivatives of Φ with respect to $\underline{x}_2$ let us evaluate first the vector

$$\varphi_l = -\tfrac{1}{2}\ \frac{\partial \varphi_l}{\partial x_{1l}} = (y_{ki}-x_{1j}A_{kij})A_{lki} \qquad (10)$$

clearly for all values of $\underline{x}_2$ $\Phi_1 = 0$ i.e.

$$(y_{ki} - x_{1j}A_{kij})\ A_{lki} = 0 \quad l=1,..,a \qquad (11)$$

Using the theorem of implicit differentiation we obtain the system

$$\frac{\partial \varphi l}{\partial x_{2m}} + \Sigma_i \frac{\partial \varphi_1}{\partial x_{1i}}\ \frac{\partial x_{1i}}{\partial x_{2m}} = 0 \quad l=1,..,a \qquad (12)$$

for every values of m=1,...,b.
Systems (12) can be solved with respect to $\partial x_{1i}/\partial x_{2m}$, which are necessary for the computation of $D\Phi/Dx_{2m}$ where D indicates total differetiation (i.e. the functional dependence of x_1 on x_2 is taken into account). Anyway we first have to compute the terms $\partial \varphi_1/\partial x_{1l}$ and $\partial \varphi_1/\partial x_{2m}$ and thus make the system (12) explicit. Differentiation of (12) gives

$$\frac{\partial \varphi_1}{\partial x_{ij}} = -\ (A_{kij}A_{lki}) \qquad (13)$$

$$\frac{\partial \varphi l}{\partial x_{2m}} = A_{lki,m}(y_{ki}-x_{1j}A_{kij}) + -A_{lki}x_{1j}A_{kij,m} \qquad (14)$$

where the differentiation with respect to a parameter is indicated by a comma followed by the parameter index. Substituting in (14) the values given by (8) for $\underline{x}_1$ we can express both (13) and (14) as functions of $\underline{x}_2$ only. Now system (12) allow us to compute the values of $\partial x_{1i}/\partial x_{2m}$ for all value $\underline{x}_2$. On the other hand the gradient of Φ is given by

$$\frac{D\,\Phi}{D\,x_{2m}} = -2\ (y_{ki}-x_{1j}A_{kij})* *(A_{kis,m}+x_{1s,m}\ A_{kis}) \qquad (15)$$

where again x_{1j} and $x_{1j,m}$ can be computed using (8) and (12). Similarly in the Gauss-Newton approximation (i.e. neglecting the terms that contain residuals with respect to those that do not) the hessian matrix is given by

$$\frac{D^2\ \Phi}{Dx_{2m}\ Dx_{2p}} = 2(x_{1j}A_{kij,p}+x_{1j,p}A_{kij})* *(A_{kis,m}+x_{1s,m}\ A_{kis}) \qquad (16)$$

A Gauss-Newton step can now be easly computed using (15) and (16) and the procedure has been completly outlined.

STATISTICAL INTERPRETATION OF RESULTS

Once convergence has been reached we are interested in investigating the goodness of the fit obtained and to this purpose the statistical properties of the parameters obtained are examined. The most important property is given by the variance-covariance matrix from which further information can be drawn. The variance-covariance matrix of $\underline{x}_2$ can be approximated in the usual way (Bard,1974)

$$V^{(x_2)} = \Phi^{-1}_{,mp}\ \Phi_{,pki}V_{kijl}\Phi_{,rjl}\ \Phi^{-1}_{,rs} \qquad (17)$$

where V is the variance-covariance matrix of data and $V^{(x_2)}$ that of parameters. V can be known or it can be estimated in addition to the parameters (Bard,1978). The indexes m, s, p and r refer to parameters and k, i, j, l to the data. The computa -

tion of $\Phi^{-1}_{,mp}$ and $\Phi_{,pki}$ from the expression is strightforward and will not be examined any further. On the other hand it is not possible to employ the same procedure to compute $V^{(x_1)}$ because $\underline{x}_1$ has not be changed independently of $\underline{x}_2$. Thus to compute $V^{(x_1)}$ defined as

$$V^{(x_1)} = E\ (\delta\underline{x}_1\ ,\ \delta\underline{x}_1^{*}) \qquad (18)$$

(where E is the expectation operator over different replicas of the data) we have to express $\delta\underline{x}_1$ as a funtion of $\delta\underline{y}$ and $\delta\underline{x}_2$. To do this let us rewrite equation (9) as follows

$$x_{1j} = B_{jik}\ y_{ik} \qquad (19)$$

where $B_{jik} = (A_{1ik}\ A_{ikj})^{-1}\ A_{1ik}$

Clearly we have

$$\begin{aligned} V^{(x_1)}_{jl} = \{ & B_{jik,m}\ y_{ik}\ E(\delta x_{2m}\ \delta x_{2r}) y_{st} B_{est,r} + \\ & + B_{jik,m}\ y_{ik}\ E(\delta x_{2m}\ \delta y_{st}) B_{est} + \\ & + B_{jik}\ E(\delta y_{ik}\ \delta x_{2m}) y_{st} B_{est,m} + \\ & + B_{jik}\ E(\delta y_{ik}\ \delta y_{st}) B_{est} \} = \\ = \{ & B_{jik,m}\ y_{ik}\ V^{(x_2)}_{mr}\ y_{st}\ B_{est,r} + \\ & - B_{jik,m}\ y_{ik}\ \Phi^{-1}_{,mp}\ \Phi_{,puv}\ V_{stuv}\ B_{est} + \\ & - B_{jik}\ V_{ikuv}\ \Phi_{,puv}\ \Phi^{-1}_{,mp}\ y_{st}\ B_{est,m} + \\ & + B_{jik}\ V_{ikst}\ B_{est} \} \end{aligned} \qquad (20)$$

Thus the computation of the statistical properties too can be carried out separately for $\underline{x}_1$ and $\underline{x}_2$ solving (17) first and then (20).
Due to the clumsiness of (20) it is often convenient to repeat the minimization procedure disregarding the the separability property of the problem and using for the parameters the values obtained in the previous minimization. Since now the set $\underline{x}_1$ is empty the whole of the statistical properties can be estimated using equation (17) only.

APPLICATIONS TO CHEMICAL KINETICS

If several reactions take place simultaneously the rate expressions are generally of the type

$$\frac{d\,c_i}{d\,t} = \sum_{j=1}^{z} k_j\ f_{ji}(\underline{c}/\underline{K}) \qquad i=1,\ldots,h \qquad (21)$$

where c_i is the concentration of the i-th reacting species and k_j are the rate coefficients or the pre exponential factor if an Arrhenius type law for the temperature dependence is made use of. The non linear parameters $\underline{K}$ are generally equilibrium costants, adsorption coefficients or activation energies. Suppose further that all concentrations are measured at time t. An approximate method due to Himmelblau et al. (1967) enables us to write

$$c_i(t_\alpha) \approx c_{i0} + \sum_{j=1}^{z} k_j \sum_{\mu=1}^{\alpha} f_{ji}(\underline{c}_\mu/\underline{K})(t_\mu - t_{\mu-1}) \qquad (22)$$

with $t_0=0$ and c_μ approximated by the averaged values measured at times $t_{\mu-1}$ and t_μ. If the linear and non linear parameters $\underline{k}$ and $\underline{K}$ are estimated by unweighted least squares we clearly obtain

$$\| \underline{y}_i - A_i\underline{k} \|^2 \qquad (23)$$

where $y_{i\mu}$ is given by $c_{i\mu} - c_{i0}$ and

$$A_{ij\alpha} = \sum_{\mu=1}^{\alpha} f_{ji}(\underline{c}_\mu/\underline{K})(t_\mu - t_{\mu-1}) \qquad (24)$$

Thus the general algorithm described in the previous paragraphs can be applied to the estimation of $\underline{k}$ and $\underline{K}$. In particular this procedure has been applied to the determination of rate coefficients of the oxidation of carbon monoxide in the presence of Hopcalite catalyst using the data and the model reported in (Brittan,1970). Rate expressions are supposed to obey the following kinetics

$$\frac{d\,p_{CO}}{d\,t} = k\ [c_1 - (2p_{O2} - p_{CO})]\ p_{CO}$$

$$\frac{d\,p_{CO2}}{d\,t} = k_1 [c_1 - (2p_{O2} - p_{CO})]\ p_{CO}$$

$$-2\ \frac{d\,p_{O2}}{d\,t} = \left\{ \frac{k_1 [c_1 - (2p_{O2} - p_{CO})]\ p_{CO}}{k_3/k_4 + p_{O2}^{\frac{1}{2}}} + k_5\ [c_2 + (2p_{O2} + p_{CO2}]\ \right\} p_{O2}^{\frac{1}{2}} \qquad (25)$$

In this expressions concentrations are replaced by partial pressures. The parameters k, k_1, k_3, k_4 and k_5 indicate rate coefficients of elementary steps, whereas c_1 and c_2 are two parameters connected with the active catalyst surface and initial concentrations. The regression has been carried out first without taking advantage of the fact that some parameters enter linearly into the least squares expression. The parameters taken in (Brittan,1970) as optimal and reported in table 1 do not correspond to a minimum of the objective function but to a point with a very large gradient.

TABLE I

k	=.00737 (min.)$^{-1}$ (mm Hg)$^{-1}$
k_1	=.00712 (min.)$^{-1}$ (mm Hg)$^{-1}$
c_1	=128.81 equiv. mm Hg
k_3/k_4	= 3.13 (mm Hg)$^{\frac{1}{2}}$
k_5	= .0085 (min.)$^{-1}$ (mm Hg)
c_2	=-323.1 equiv. mm Hg

Consequently the steps computed by the algorithm are very small and this deceives the program user into thinking that convergence has been attained. Different starting points give rise to the location of a large number of local minima, all of them with similar values of the objective function. The situation is entirely different if separability property is made use of. In fact in this case the regression analysis is carried out with respect to the two non linear parameters c_1 and k_3/k_4 so that a true minimum is always achieved after few iterations. Further the global minimum can be located using a small number of different starting points. In the case examined only two different

local minima have been detected with values of the variance very close to each other.
The values of c_1 and k_3/k_4 are indicated in table II.

TABLE II

Parameters	1st minimum	2nd minimum
c_1	125.4	128.59
k_3/k_4	-.0002432	-5.551
Δ (c_1)	±7.2	±7.25
Δ (k_3/k_4)	$\pm 10^{-6}$	±.01

Thus minima of the objective function are meaningless due to the presence of a negative ratio k_3/k_4 with small standard deviations. If a bound for the lower limit of k_3/k_4 is introduced, it always turns out to be a binding constraint, so that little or no information can be drawn from it. As a consequence of these results we are induced to reject the proposed model for lack of significance of the parameters estimated, the values proposed being the outcome of the defective regression analysis.

CONCLUSION

In this paper we have described an algorithm to carry out regression analysis in the multiresponse case only with respect to those parameters that do not enter linearly into objective function.
In addition to computational efficiency the algorithm posseses increased robustness, improved reliability and the capability of doing away with many of the local minima encountered if the separability property is not taken advantage of.

NOMENCLATURE

A = see equations (1) and (2)
a = number of linear parameters
B = see equation (19)
b = number of non linear parameters
c_i = concentrations of i-th species
E = expectation operator
f = see equation (21)
h = number of concentrations measured
K = non linear parameters in rate expressions
k = linear rate coefficients
n = (a+b)
p = partial pressures
t = times
V = variance-covariance matrix of data
$V(x_1)$ = variance-covariance matrix of linear parameters
$V(x_2)$ = variance-covariance matrix of non linear parameters
x_1,q = linear parameters
x_2,c = non linear parameters
y = data

Greek

Δ = standard deviation
δ = small change
Φ = objective function
φ = see equation (10)

Superscripts

+ = indicates pseudoinverse of a matrix
* = indicates transpose of a matrix
' = indicates error effective measurements

Indexes

CO
CO_2
i
j
k
l
m
O_2
p
r
s
t
u
v
α
μ

REFERENCES

Bard, Y.(1974). Non linear parameter estimation, Academic Press, New York

Brittan,M.I., H. Bliss and C.A. Walker. Kinetics of the hopcalite-catalyzed oxidation of carbon monoxide.AIChE J., 16, 2, 305- 314.

Doví, V.G., E.P. Arato (1985). Reduction of correlation factors in the estimates of catalytic kinetics parameters. V Congresso Nazionale di Catalisi. Padova (ITALY).

Golub, G.H. and V. Pereyra (1973). The differentiation of pseudo-inverses and non linear least square problems whose variable separate. SIAM J. Numer. Anal., 10, 2, 413-432.

Golub, G.H. and V. Pereyra(1976). Dif ferentiation of Pseudo-inverses, Separable Nonlinear least square problems and other tables. (in: Generalized Inverses and Applications), Academic Press, New York.

Himmelblau, D.M., C.R. Jones and K.B. Bischoff (1967). Determination of Rate Constants for Complex Kinetics Models. I & E Chem. Fund., 539-542.

Provencher, S.W., J. Hendrix, L.Maeyer and Paulussen (1978). Direct analysis of continuous relexation spectra. J. Chem. Phys., 69, 4273-4276.

Provencher, S.W. (1979). Private comunication.

Ruhe, A., P.A. Wedin (1980). Algorithms for separable non linear least squares problems. SIAM Review, 22, 3, 318-337.